

proyecto: ID2016/110

INFORME FINAL (junio 2017)

Proyectos de innovación docente de la USAL

Referencia del proyecto: ID2016/110

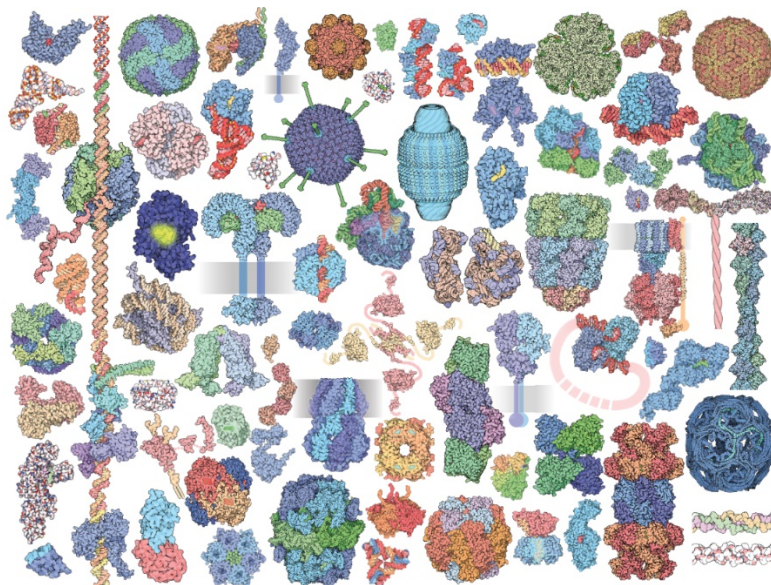
TÍTULO: Aplicación de recursos informáticos y bioinformáticos al aprendizaje de la Bioquímica (autoaprendizaje y aprendizaje a lo largo de la vida).

AUTORES: Juana Gutiérrez de Diego y F. David Rodríguez García

Departamento de Bioquímica y Biología Molecular

Facultad de Ciencias Químicas

Universidad de Salamanca



Protein Data Bank (www.rcsb.org)

INTRODUCCIÓN

En el presente proyecto nos propusimos involucrar a los estudiantes en un trabajo que pretende abordar el problema de la existencia de información muy abundante, inabarcable en muchos casos, de una forma racional y ordenada (Luo, J. 2014). Utilizamos el ordenador conectado a la red como infraestructura principal para acercarnos a la información, analizarla, cribarla y ordenarla. El ordenador, o dispositivos similares, conectados a la red, son hoy un recurso imprescindible en la labor de los científicos y profesionales de diferentes ámbitos. (Whitehead, J.P & Pence, H.E., 2002; Brown, J.A.L. 2016).

Nuestro objetivo se centró en el empleo de bases de datos “en línea” relevantes en nuestro campo, como se señala más adelante. Las bases de datos que proponemos no solo constituyen una **fuentes primaria** y fiable de información, sino que, además, proporcionan recursos de trabajo muy útiles como, por ejemplo, la visualización dinámica de estructuras moleculares tridimensionales complejas, el cálculo rápido de parámetros fisicoquímicos, la visualización de fenómenos celulares a escala real, la compilación y análisis de datos a gran escala o el uso de metadatos, etc.

Nuestra propuesta no pretendía que los estudiantes alcanzaran un conocimiento exhaustivo y pormenorizado de los medios de trabajo y de los contenidos abordados. Las bases de datos objeto de este proyecto son plataformas complejas que ofrecen información muy voluminosa y muchas y variadas herramientas. Buscamos un acercamiento de los estudiantes al manejo de la información, la optimización del tiempo y la obtención de resultados valiosos en tiempos relativamente cortos. Además, y debido a los cambios tecnológicos continuos, hay un elemento de adaptabilidad en el proceso de aprendizaje, esencial para enfrentarse a los desafíos del futuro profesional, sea cual sea el ámbito de trabajo desempeñado. Tratamos de abrir una puerta e indicar algunas rutas esenciales que facilitaran a los estudiantes la incorporación de la bioinformática en sus esquemas de aprendizaje y desarrollo profesional (autoaprendizaje y aprendizaje a lo largo de toda la vida).

Las actividades organizadas sirven a los objetivos siguientes:

- Comprensión de estructuras moleculares tridimensionales complejas
- Observación de la animación de fenómenos moleculares: reacciones y rutas metabólicas, interacciones específicas, motores moleculares...
- Análisis de secuencias de proteínas y ácidos nucleicos: comparaciones, localización de dianas de modificaciones con repercusión funcional, etc.
- Recopilación de datos relevantes acerca de una proteína determinada: propiedades físico-químicas, índices hidropáticos, etc.
- Empleo de herramientas de organización bibliográfica (aplicable a las tutorías de los TFG).

HERRAMIENTAS INFORMÁTICAS

Las bases de datos y procedimientos informáticos que empleados en las asignaturas de Grado se describen brevemente a continuación:

-**PDB**: <http://www.rcsb.org/pdb/home/home.do>

El banco de datos de proteínas (*The Protein Data Bank*, PDB) es un repositorio de estructuras tridimensionales de macromoléculas, desde proteínas de pequeño tamaño a complejas maquinarias moleculares como el ribosoma. PDB se actualiza semanalmente.

-ExPASy: <https://www.expasy.org/>

ExPASy es un portal bioinformático que da acceso a bases de datos de carácter científico y técnicas aplicables en diferentes campos de las ciencias de la vida (proteómica, genómica, filogenia, biología de sistemas, genética de poblaciones, metabolómica, enzimología, transcriptómica, etc).

(ver páginas de inicio para ambas plataformas en el Anexo Ia).

-PUBMED: <http://www.ncbi.nlm.nih.gov/pubmed>

PubMed incluye más de 26 millones de citas relacionadas con la literatura biomédica, ciencias de la vida, ciencias químicas, bioingeniería, etc. *PubMed* también da acceso a otros recursos relevantes en el campo de la Bioquímica y la Biología Molecular. Esta plataforma depende del Centro Nacional para la Información Biotecnológica (NCBI, USA).

-RefWorks & Write and Cite: <https://www.refworks.com/es/>

RefWorks está concebido para ayudar a recopilar, gestionar, almacenar y compartir información, así como para insertar citas y redactar bibliografías en más de 3000 estilos de bibliográficos.

Write-N-Cite es una prestación de RefWorks que permite generar citas y bibliografías desde un documento de Word.

(ver páginas de inicio para ambas plataformas en el Anexo Ib.)

ORGANIZACIÓN. SESIONES DE TRABAJO

Los profesores elaboraron los materiales de trabajo y establecieron los criterios de evaluación.

I. Actividades dirigidas a los estudiantes de Grado en Química

En una tutoría grupal de corta duración se hizo una breve introducción a los medios informáticos; no tanto a la mecánica, cuanto a la importancia del manejo y compilación de información, a las estrategias de análisis, la detección de fuentes primarias y secundarias, fiabilidad de los sitios web, etc. Las sesiones tuvieron una duración de 50 minutos. Durante las mismas, los alumnos pudieron consultar a los profesores acerca de sus dificultades y dudas.

Sesión PDB.

Esta sesión se dedicó al estudio monográfico de una proteína, incluyendo aspectos estructurales y funcionales de la misma (ver el guion en el anexo II).

Sesión ExPASy.

Seguimos el mismo patrón que el dedicado al portal PDB. Estudiamos las propiedades básicas de una enzima mediante el acceso a la página ENZYME (ver el guion anexo III). Empleamos diferentes herramientas del portal para comparar secuencias, calcular parámetros físico-químicos y realizar traducciones de secuencias de DNA a las proteínas codificadas (ProtParam, TRANSLATE, BLAST ...)

EVALUACIÓN

Ambas actividades fueron evaluadas por los profesores mediante una rúbrica que recogía los aspectos más relevantes de la actividad, basada en la cumplimentación correcta de los datos y cálculos solicitados en el guion. Las dos actividades se

incluyeron en el programa de evaluación continua de la asignatura correspondiente (Ver anexo IV).

Resultados

1. Participación

La tabla 1 muestra el número de alumnos que participaron en ambas sesiones de trabajo:

Tabla 1. Alumnos que se presentaron a las actividades, PDB y EXPASy

Actividad	Alumnos presentados	Alumnos no presentados
PDB	75	6
EXPASy	71	10

2. Calificaciones

Los informes presentados fueron calificados con una nota máxima de 1 punto. La nota media de los presentados fue de **0,91 \pm 0.07** para la actividad PDB y de **0,9 \pm 0.06** (media \pm SEM).

3. Encuesta a los estudiantes

Elaboramos una encuesta que los alumnos respondieron en la plataforma *Studium*. Los resultados más relevantes se presentan a continuación.

Respuestas contabilizadas: **52** (aproximadamente, un 70% de los alumnos que realizaron las actividades)

Preguntas:

-Valora globalmente las actividades desarrolladas en el aula de informática de 1 (muy negativo) a 5 (muy positivo)

Sesión en el aula de informática sobre la base de datos <i>Protein Data Bank</i> (PDB)	3.2
Sesión en el aula de informática sobre la base de datos EXPASy	3.2

-Valora de 1 (no estás de acuerdo en absoluto) a 5 (estás completamente de acuerdo), los siguientes aspectos relacionados con las actividades desarrolladas del aula de informática:

1. En estas actividades he aprendido bastante sobre el manejo de las bases de datos visitadas	3.3
2. Los profesores me han atendido adecuadamente durante las sesiones	4.6
3. Los guiones proporcionados son claros	3.8
4. Considero que las actividades desarrolladas me han resultado difíciles	2.6

-Comentarios anónimos seleccionados

-Escribe sugerencias acerca de las actividades desarrolladas en el aula de informática: organización, dificultad, tiempo dedicado, acceso a otros recursos (¿cuáles?), etc.

Las sesiones de ordenador, creo que deberían ser de más tiempo, porque si no realmente no se entiende muy bien lo que se está haciendo con las bases de datos.

Las actividades en el aula de informática también han estado muy bien ya que hemos aprendido donde buscar información en las distintas páginas web, no ha sido difícil de asimilar, pero quizás al principio estaba un poco perdido con las bases de datos, pero eso se solventó con la gran atención por parte de los profesores.

La organización es correcta. El nivel de dificultad me parece adecuado para una asignatura de tercero de carrera. El tiempo dedicado para manejar una base de datos... pues desde mi punto de vista es muy poco, sería mejor dedicarle dos horas en dos días seguidos o incluso el mismo día.

Las páginas de acceso nos proporcionaban mucha información, pero se hacía algo difícil al no estar familiarizados con ese tipo de páginas.

El tiempo para desarrollar las dos actividades realizadas en el aula de informática ha sido el necesario para poderlas hacer enteras. Ninguna de las dos fueron actividades complicadas, la primera tenía algo más de dificultad, aunque una de las razones pudo ser porque era la primera vez que visitábamos esas bases de datos y no teníamos soltura.

En una hora poco se puede hacer, pero me parecen entretenidas y didácticas, además cualquier duda que tengamos nos las solucionáis bien.

Creo que se deberían omitir estas actividades, son poco provechosas

Estas eran las más difíciles porque había que completarlas en directo (no las podíamos mirar antes en casa) además que teníamos que usar páginas complejas con muchas pestañas. Pero ayudándonos entre varios y estando atento se sacaban adelante.

En la primera práctica el tiempo fue un gran impedimento ya que la mayoría de textos están en inglés y traducir y entender lleva su tiempo. Aun así, permite estar atento y sin despistarte para que dé tiempo terminar la actividad

Creo que con un número más reducido de alumnos se desarrollaría mejor, porque nunca hemos manejado esas bases de datos y tenemos muchas dudas.

Estaba bien organizado, la dificultad no ha sido mucha ya que al estar en inglés se entendía bien, el tiempo dedicado ha sido el justo ya que para hacer el seminario o el minipóster hemos recurrido allí

CONSIDERACIONES DE LOS PROFESORES

Estamos convencidos de que esta actividad es relevante para la formación de los graduandos ya que pretende proporcionar un marco de aprendizaje útil, aplicable al futuro profesional. Aun cuando en términos generales la actividad tiene una valoración aceptable por parte de los alumnos que han respondido a la encuesta, es también obvio que se apuntan ciertas dificultades relacionadas con el tiempo necesario para la realización de la sesión, el desconocimiento previo de estas bases

de datos complejas y el hecho de que la lengua de las mismas sea el inglés. Algunos alumnos, por su parte, apuntan la inutilidad de esta actividad.

La valoración de los estudiantes es para nosotros muy importante y nos sirve para la reflexión y para introducir algunas modificaciones metodológicas que puedan servir de manera más eficaz a la organización y aprovechamiento de la sesión. En cuanto a los comentarios relacionados con el empleo de la lengua inglesa en estas bases de datos, no podemos hacer sino intentar convencer a los estudiantes de que el conocimiento del inglés es imprescindible para su carrera profesional.

Por otra parte, muchos alumnos consideran que la actividad no presenta mucha dificultad y, además, ha sido resuelta en el tiempo establecido y con una calificación media muy alta. Recogemos la sugerencia referida a la familiarización previa con estas páginas, ampliando las instrucciones mediante un documento escrito proporcionados los estudiantes con antelación.

II. Actividades dirigidas a los estudiantes que realizan el TFG

En una tutoría inicial establecimos los criterios de trabajo, calendario, compromiso, etc. y planteamos un espacio monográfico dedicado al acceso y manejo de la base de datos *PubMed* y al empleo de una herramienta de organización bibliográfica (*RefWorks* y *Write and cite*). Además, se resolvieron dudas mediante el correo electrónico y en las tutorías presenciales programadas a lo largo del período de supervisión del TFG. Este procedimiento se aplicó al TFG tutorizado por los profesores que firman esta memoria, titulado *Acuaporinas*, que será presentado el próximo mes de julio de 2017.

Evaluación

De acuerdo a la normativa, elaboramos un informe breve de valoración cualitativa del TFG en su conjunto, que se adjunta a la documentación que recibe el tribunal correspondiente.

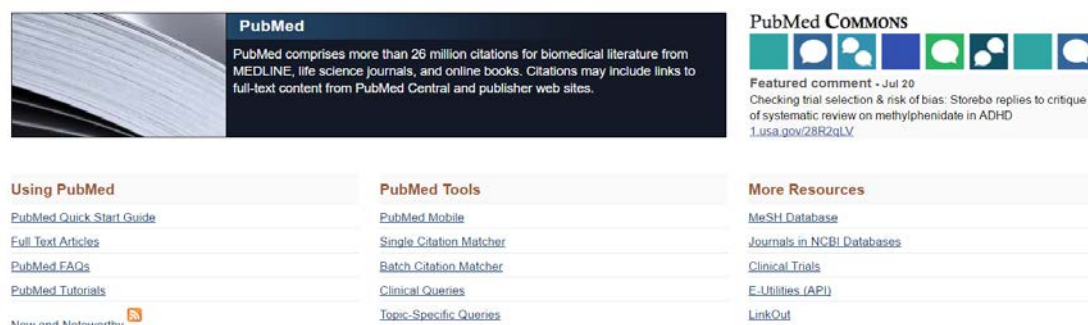
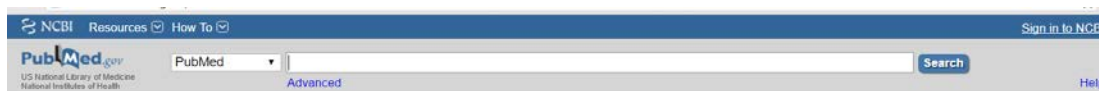
CONSIDERACIONES DE LOS PROFESORES

En nuestra experiencia, es llamativo que los estudiantes que comienzan su TFG no tengan un conocimiento básico del acceso y manejo de las bases de datos bibliográficas primarias, así como de gestores bibliográficos que facilitan mucho el trabajo. Por ello, seguiremos dedicando un tiempo específico, en el marco de la acción tutorial, al objetivo marcado que facilite a los estudiantes el empleo racional, adecuado y efectivo de herramientas imprescindibles para su carrera profesional.

REFERENCIAS

- Luo, J. (2014) Teaching the ABCs of bioinformatics: a brief introduction to the Applied Bioinformatics Course. *Briefings in Bioinformatics*, 15 (6), pp 1004 -1013.
- Brown, J.A.L. (2016). Evaluating the Effectiveness of a Practical Inquiry-Based Learning Bioinformatics Module on Undergraduate Student Engagement and Applied Skills. *Biochemistry and Molecular Biology Education*, 44 (3), pp 304–313.
- Whitehead, J. P. & Pence, H. E. (2002) Multimedia in Biochemistry and Molecular Biology Education Using Computers to Teach Biochemistry. *Biochemistry and Molecular Biology Education*, 30 (3), pp. 206–207.

ANEXO I. Páginas de Inicio de las bases de datos (acceso en 2017)



ANEXO II

Base de datos del PDB (*Protein Data Bank*)

Nombre y Apellidos:

En un laboratorio de investigación se ha aislado y purificado una proteína a partir de *Arabidopsis thaliana*. Los investigadores han depositado información de dicha proteína en la base de datos *Protein Data Bank* (www.rcsb.org/pdb) con el código de acceso **5KSD**.

Responde a las siguientes preguntas a partir de la información que la base de datos ofrece y de la separata de Pedersen y cols 2007 que encontrarás en *Stodium*.

Nota: Antes de responder debes familiarizarte con el portal PDB, incluyendo los accesos y conexiones a otras bases de datos.

¿Qué tipo de organismo es *Arabidopsis thaliana*?

Datos generales de la proteína

Nombre:

¿Cuál es su peso molecular?

¿De cuántos residuos consta la proteína cristalizada?

2. Acerca de la estructura de la cadena polipeptídica:

¿Cuántos dominios citoplásmicos y transmembrana tiene?

Describe brevemente su estructura secundaria

Indica qué residuos ocupan las posiciones 12, 20, 55, 300 y 841:

¿Qué estructura secundaria adoptan los residuos en las posiciones 235-254

¿Qué residuos, y en qué posiciones conforman una hélice π ?

Señala una región sin estructura secundaria definida que comprenda más de 5 residuos. Indica entre qué posiciones

Escribe la carga neta, *in situ*, del tetrapéptido del extremo amino-terminal a pH=7:

¿Podríamos encontrar puentes disulfuro en esta proteína? Establece la posición de dos hipotéticos puentes disulfuro

3. Acerca de su función.

Función/funciones de esta proteína:

¿Cuál es el significado de **P** en las ATPasas tipo P

Escribe dos ejemplos de ATPasas tipo P

Representa un esquema con la estructura de la bomba de protones e indica la localización del dominio de unión a nucleótidos (N), y del dominio donde tiene lugar la fosforilación de la proteína (P). (ver figura 1 de la publicación, doi:10.1038/nature06417)

La cavidad por la que entran los protones en la proteína está flanqueada por dos aminoácidos, ¿cuáles? (ver figura 3 y 4 de la publicación, doi: 10.1038/nature 06417).

¿Qué dirección siguen los protones?

ANEXO III

Base de Datos EXPASY

Nombre y Apellidos:

Accede en la página EXPASY (www.expasy.org) a la base de datos ENZYME con el código de acceso EC **3.1.1.84** Responde a las siguientes preguntas a partir de la información que la base de datos ofrece sobre la enzima.

Nota: Antes de responder debes familiarizarse con el portal EXPASY, incluyendo los accesos y conexiones a otras bases de datos.

- Nombre (s) de la proteína:
 - Significado de los dígitos del código EC:
 - Escribe la reacción que cataliza:
 - Escribe el número **EC** de los enzimas: Acetilcolinesterasa y Clorofilasa
- Indica qué relación tienen con la proteína **3.1.1.84**

Mediante la consulta de las herramientas indicadas a continuación, responde a las preguntas formuladas:

Base de Datos BRENDA (a la que accedes desde ENZYME). Busca los datos para ***H. sapiens***.

- Escribe la estructura y nombre de un inhibidor de la enzima:
- Indica los valores de K_m para cocaína (enzima no modificada) y heroína en la especie *Homo sapiens*

En la Base de datos UNIPROT: <http://www.uniprot.org/uniprot/O00748>. Copia la secuencia de aminoácidos de la isoforma 1 (copia el formato **FASTA sin el encabezamiento identificativo**) para posteriores análisis. Utiliza las herramientas de Expasy que se indican a continuación (en orden alfabético en *Resources A-Z*) para responder a las preguntas.

GOR (protein secondary structure prediction)

- Determina qué porcentaje de aminoácidos forma hélices alfa y señala la secuencia de la primera hélice alfa formada por 5 residuos ¿Cuántas hélices 3_{10} aparecen en esta predicción?

PROTPARAM.

- Determina el punto isoeléctrico (pI) teórico y el Peso molecular de esta proteína. ¿Cuál es su aminoácido más abundante?

PEPTIDE CUTTER

- ¿Cuántas rupturas se obtienen tras la digestión de la proteína con Tripsina? ¿Qué característica común comparten los enlaces hidrolizados por Tripsina?
- ¿Cuántas rupturas se obtienen con Enteroquinasa? ¿Y con Trombina?

TRANSLATE

A partir de la secuencia del siguiente fragmento del DNA inicial del gen 5' ATGACTGCTCAGTCCCGCTCT 3', determina la secuencia del péptido que codifica, su peso molecular, el punto isoeléctrico (pI) teórico y carga neta a pH=1.

ANEXO IV Evaluación de la asignatura Bioquímica (Grado en Química)

